

*Citation for published version:*

Yin, S, Sui, J, Chiu, Y-C, Chen, A & Egnér, T 2019, 'Automatic prioritization of self-referential stimuli in working memory', *Psychological Science*, vol. 30, no. 3, pp. 415-423. <https://doi.org/10.1177/0956797618818483>

*DOI:*

[10.1177/0956797618818483](https://doi.org/10.1177/0956797618818483)

*Publication date:*

2019

*Document Version*

Peer reviewed version

[Link to publication](#)

Yin, Shouhang ; Sui, Jie ; Chiu, Yu-Chin ; Chen, Antao ; Egnér, Tobias . / Automatic prioritization of self-referential stimuli in working memory. In: *Psychological Science*. 2018, Vol. 30, no. 3, pp. 415-423. (C) The Authors, 2019. Reprinted by permission of SAGE Publications.

**University of Bath**

**Alternative formats**

If you require this document in an alternative format, please contact:  
[openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# **Automatic prioritization of self-referential stimuli in working memory**

Shouhang Yin<sup>a</sup>, Jie Sui<sup>b</sup>, Yu-Chin Chiu<sup>c</sup>, Antao Chen<sup>a\*</sup>, Tobias Egner<sup>c</sup>

<sup>a</sup>Key laboratory of Cognition and Personality of the Ministry of Education, Faculty of Psychology, Southwest University, Chongqing, 400715, China

<sup>b</sup>Department of Psychology, University of Bath, UK

<sup>c</sup>Center for Cognitive Neuroscience, and Department of Psychology and Neuroscience, Duke University, Durham, NC, USA

\*Correspondence should be addressed to: Antao Chen, Faculty of Psychology, Southwest University, BeiBei District, Chongqing, China. E-mail: xscat@swu.edu.cn

## **Abstract**

People preferentially attend to external stimuli that are related to themselves as compared to others. Whether a similar self-reference bias applies to internal representations, maintained in working memory (WM), is presently unknown. We tested this possibility in four experiments, where participants were first trained to associate social labels (self, friend, stranger) with arbitrary colors, and then performed a delayed match-to-sample spatial WM task on color locations. Participants consistently responded fastest to WM probes at locations of self-associated colors (Experiments 1-4). This self-bias was not driven by differential exogenous attention during encoding or retrieval (Experiments 1 and 2), but by internal attentional prioritization of self-related representations during WM maintenance (Experiment 3). Moreover, self-prioritization in WM was non-strategic, as this bias persisted even under conditions where it hurt WM performance. These findings document an automatic prioritization of self-referential items in WM, which may form the basis of some egocentric biases in decision-making.

**Key words:** self-reference; self-prioritization effect; self-bias; working memory; internal attention

## Introduction

A large body of research has shown that people are biased towards preferentially processing self-related information compared with other social information (Banaji & Prentice, 1994): self-referential stimuli capture exogenous attention, resulting, for instance, in faster responses to one's own face than other faces (Alexopoulos, Muller, Ric, & Marendaz, 2012; Keyes & Brady, 2010; Liu, He, Rotstein, & Sui, 2016), and one's own face or name can act as a potent task-irrelevant distracter (Brédart, Delchambre, & Laureys, 2006; Devue & Brédart, 2008; Harris & Pashler, 2004). Moreover, stimuli that are encoded in a self-referential way are remembered better than those that are not (Kesebir & Oishi, 2010; Symons & Johnson, 1997). While self-reference advantages in exogenous attention and long-term memory are thus well established, it is presently not known whether self-reference also modulates the prioritization of information in *working memory* (WM), which allows us to temporarily keep information in mind for additional cognitive processing and the guidance of actions (Baddeley, 2003).

Many current views posit that information is maintained in WM via internally directed attention (Awh & Jonides, 2001; Barroulliet, Bernadin, & Camos, 2004; Oberauer, 2009; Chun, Golomb, & Turk-Browne, 2011; D'Esposito & Postle, 2015; Gazzaley & Nobre, 2012; Kiyonaga & Egnér, 2013). Considering the pull that self-referential stimuli in the outside world exert on exogenously directed attention, it is plausible that self-referential representations in WM would similarly attract internal attention, and thus be prioritized over other, not self-related items held in WM. This is an important topic to study not only for understanding mechanisms of WM and self-referential processing, but also because maintaining and evaluating information in WM is central to decision-making and cognitive control (Baddeley, 2003; D'Esposito and Postle, 2015). Thus, a strong bias for representations of self-referential stimuli could have undesirable social side effects, such as promoting egocentric decisions. In the present study we therefore aimed to test, for the first time, whether people exhibit a self-bias in WM.

A large literature on WM retro-cuing effects has shown that the differential prioritization of items in WM can be measured by the relative speed and accuracy with which these items are accessed upon cued retrieval (Griffin & Nobre, 2003; Souza & Oberauer, 2016). We thus here tested whether this holds true for self-related stimuli in WM. Specifically, we addressed this question over 4 experiments by adapting as a spatial WM task a recently developed protocol where newly formed associations between the self and arbitrary stimuli modulated perceptual processing (Sui, He, & Humphreys, 2012). This approach avoids the confounding impact of familiarity on self-reference effects (Sui et al., 2012, Sui, Sun, Peng, & Humphreys, 2014; Humphreys & Sui, 2015; Sui & Humphreys, 2015) and thus allowed us to probe self-prioritization in WM in a tightly controlled manner.

Each experiment involved two tasks: the first was an associative learning task that served to create links between arbitrary stimuli (here: colors) and social associations (self, friend and stranger) (cf. Sui et al., 2012). The second was a delayed match-to-sample spatial WM task that assessed the potential impact of these self- vs. other-related color associations on their prioritization in WM. To preview the results, we found that self-related items are reliably prioritized in WM, that this self-bias occurs specifically during WM maintenance, and that it appears to be obligatory, as it is present even when it hurts performance.

## **Experiment 1**

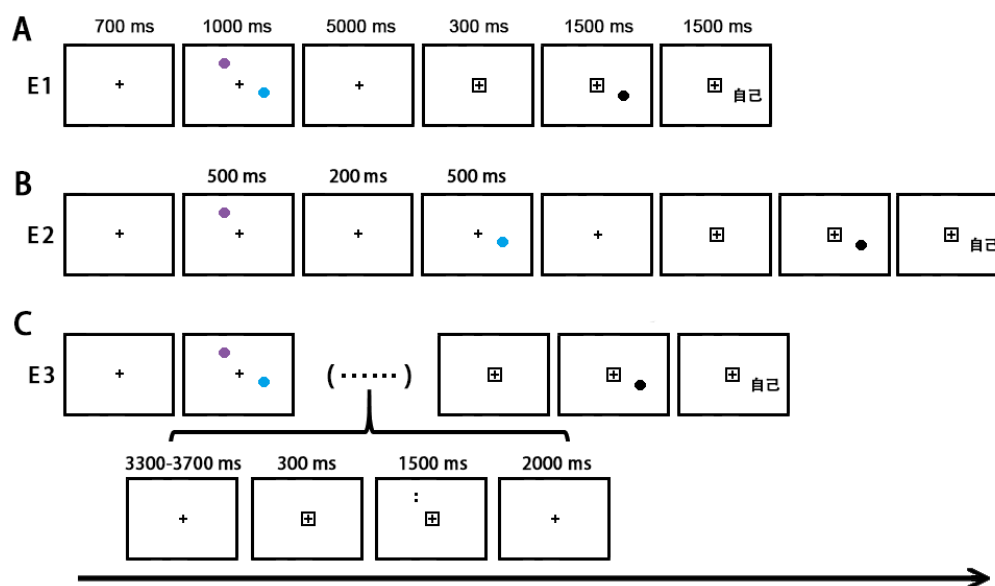
### ***Method***

***Participants.*** We used GPower (Faul, Erdfelder, Lang, & Buchner, 2007) to estimate the effect size from a comparable previous paper (Sui et al., 2012), which was 1.42. With a desired power of 0.8 for detecting an effect at an alpha level of 0.05, we would need a minimum sample size of 18. Thus, we set a target sample size of  $N = 25$  for all experiments. Due to variability in participant scheduling reliability, we ended up with  $N = 27$  in Experiments 1 and 2,  $N = 25$  in Experiment 3, and  $N = 23$  in Experiments 4. There was no overlap in participants across experiments, and all participants in all experiments reported normal color vision and normal or corrected-to-normal acuity. This study was approved by the University Human Ethics Committee of Southwest University (China). Twenty-seven undergraduates from Southwest University (7 males; mean age = 22.47 years, range = 18–23) completed Experiment 1.

***Stimuli and Procedure.*** All four experiments were run on a PC and an 18.5-in. monitor ( $1,366 \times 768$  at 60 Hz), using E-prime software (Version 2.0). The four experiments each contained two tasks, with the first (an associative learning task) being the same across experiments. The associative learning task was a modified version of the task used in Sui et al.'s study (2012), where participants were trained to form associations between arbitrary colors and social labels (self, friend, stranger) (see the Supplemental Material for procedure details).

The second task was a delayed match-to-sample spatial WM task and was performed immediately after the learning phase. On each trial, two solid colored circles (blue, green and purple,  $1.2^\circ \times 1.2^\circ$ ) were presented at nonadjacent two of eight possible vertices of an invisible regular octagon ( $9.6^\circ \times 9.6^\circ$ ). As shown in **Fig. 1A**, a trial started with a 700-ms fixation cross that remained on screen throughout the trial, followed by two colored circles shown for 1000 ms. Participants were asked to remember the locations and social associations linked to these colors (based on the prior learning task). Then, the trial entered a 5000 ms maintenance stage, after which a square was displayed around the fixation cross for 300 ms, which signaled the forthcoming presentation of a probe target and encouraged central fixation prior to probe onset. A target (a black circle) was then presented for 1500 ms at one of the eight possible locations, during which time participants had to judge whether the

target location matched either of the two remembered locations, using the index and middle fingers of right hand on the keypad keys “1” and “2”. The assignment of response finger to “match” or “non-match” responses was counterbalanced across participants. Target presentation was terminated by the key press or after 1500 ms. If the target matched either of the two remembered locations, a label word was presented at the target location for 1500 ms, and participants were required to judge whether the label word matched the remembered color in this location; on non-match trials, the trial ended with a 1500-ms blank screen instead. The goal of this color-label probe was to ensure that participants actively remembered the label words that the colors were associated with.



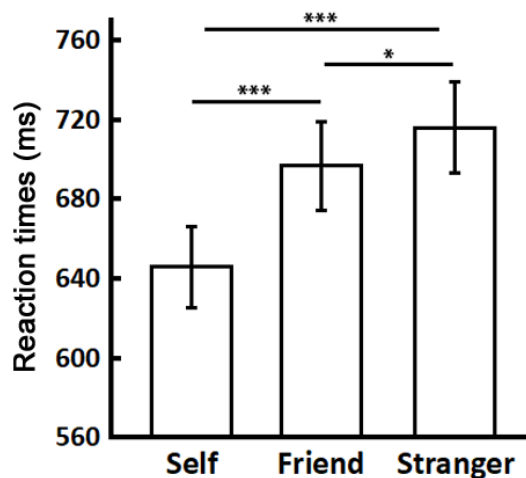
**Fig. 1** Trial timing and example stimulus displays in the spatial WM tasks of Experiments 1 (A), 2 (B), and 3 (C), respectively. The Chinese characters “自己” are identical to the English word “self”.

The different possible combinations of the color memory items resulted in three trial types or pairings: Self-Friend, Self-Stranger and Friend-Stranger. Each of these trial types occurred 96 times, including 24 match trials for each of the two items and 48 non-match trials. Altogether, there were 288 trials, including 48 self-matched trials, 48 friend-matched trials, 48 stranger-matched trials and 144 non-match trials, broken down into 6 blocks of 48 trials each (following 10 practice trials); trial types were presented in random order. Sorted by the social associations of match probes, this experiment was a 3-level single-factor (social association: Self, Friend, and Stranger) within-subjects design. The experiment lasted around 80 minutes.

The experiments reported here were not formally pre-registered. All materials and data are available on the Open Science Framework website (<https://osf.io/k2vp7/>).

## Results

For all experiments, only correct responses with RTs above 200 ms and within 2.5 standard deviations (SDs) from the subject-specific mean (for each condition) were used for the RT analysis, eliminating less than 1% of trials overall. Mean RT was used for all analyses. The associative learning task data successfully replicated those of Sui et al.'s (2012), documenting a successful learning of color-word associations with a significant bias for self-related items (results for the associative learning phase from all experiments are shown in the Supplemental Material). In the spatial WM task, the mean accuracies of the probe and label response were 98% and 97% respectively. Since every subject's accuracy was higher than 95%, we did not analyze the accuracy data further. (The same applies for Experiments 2-4). A repeated-measures one-way analysis of variance (ANOVA) on RT data of location probe match trials showed a significant main effect of social association ( $F(2, 52) = 28.70, p < 0.001, \eta^2 = 0.53$ , see **Fig. 2**), with faster responses to Self-match trials than to Friend-match trials ( $p < 0.001$ ) and to Stranger-match trials ( $p < 0.001$ ), as well as faster responses for Friend-match trials than Stranger-match trials ( $p < 0.05$ ). For the RT data of the color-label probe (following location-match responses), the results in this and the other experiments were similar to the results of associative learning task, documenting a robust self-bias (see Supplemental Material for color-label probe results of all experiments).



**Fig. 2** Mean RTs ( $\pm$  SEM) for Self-match, Friend-match and Stranger-match conditions of the location probe of spatial WM task in Experiment 1. \*\*\* represents  $p < 0.001$ , \* represents  $p < 0.05$ .

## Discussion

The results of the WM task document a significant self-prioritization effect, with faster retrieval for self-referential items. Since the target location probe consisted of a neutral black circle, this result cannot reflect external attentional capture by a self-referential stimulus during WM retrieval. Rather, the effect must stem from biased processing during the encoding or maintenance stage. Notably, at encoding, the two color items were presented simultaneously, thus directly competing for attentional

resources (Desimone & Duncan, 1995), such that the self-bias observed in Experiment 1 could in theory reflect a consequence of greater bottom-up capture of attention to the self-associated color during encoding. Similarly, previous studies have shown serial position effects of item encoding in WM, and the self-bias may result from encoding the self-referential item first (Blalock & Clegg, 2010; Hay, Smyth, Hitch, & Horton, 2007). In Experiment 2, we therefore tested whether the WM self-bias is driven by the attentional (and/or serial) prioritization of self-referential stimuli during the encoding stage, by presenting the color cues serially.

## ***Experiment 2***

### ***Method***

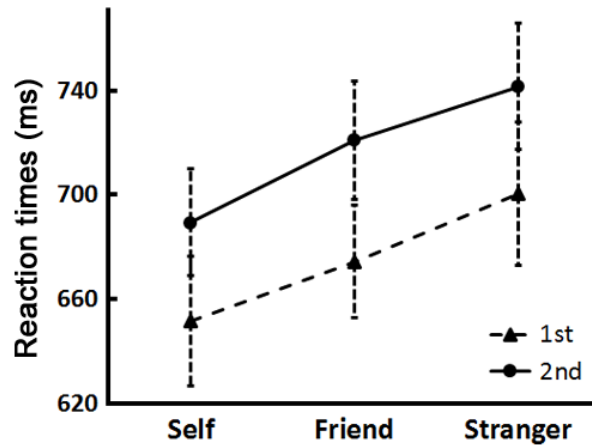
**Participants.** Twenty-seven undergraduates from Southwest University (9 males; mean age = 21.06 years, range = 18–23) completed the experiment.

**Stimuli and Procedure.** The stimuli and procedure were identical to those in Experiment 1, except that we modified the item presentation mode in the WM task to sequential presentation. The two WM color items were presented one by one with an interval of 200 ms between them (see **Fig. 1B**). The task involved two within-subject factors: serial order (2 levels: 1st vs. 2nd) and social association (3 levels: Self, Friend and Stranger). The different trial types were presented 144 times each. For each of these, there were 72 trials where one of the two items was presented first, and 72 trials where the other one was presented first. At the probe stage, each of these conditions was associated with 36 match and 36 non-match trials. The task was broken down into 6 blocks of 72 trials (following 10 practice trials). The experiment lasted around 100 minutes.

### ***Results***

The learning task and color-label probe data replicated the results of Experiment 1, documenting robust learning and maintenance of color-label associations (see Supplemental Material). Mean WM accuracy was 98% for the location and 96% for the color-label probe. A repeated-measures 2 (order: 1st vs. 2nd)  $\times$  3 (social association: Self, Friend and Stranger) ANOVA on RT data of location probe match trials revealed that the two main effects were significant ( $F(1, 26) = 11.02, p < 0.01, \eta^2 = 0.30$  for order;  $F(2, 52) = 11.46, p < 0.01, \eta^2 = 0.31$  for social association). Critically, there was no interaction between serial order and social association ( $F(2, 52) = 0.42, p = 0.66, \eta^2 = 0.02$ ). Regardless of presentation order, the mean RT of Self-match trials was significantly shorter than Friend-match ( $p < 0.01$ ) and Stranger-match trials ( $p < 0.01$ ), as well as for Friend-match compared to Stranger-match trials ( $p < 0.01$ ). **Fig. 3** displays mean RTs for all conditions.





**Fig. 3** Mean RTs ( $\pm$  SEM) for Self-match, Friend-match and Stranger-match conditions of the location probe of spatial WM task in Experiment 2. The lower line represents the results when matched item is presented 1st and the upper line represents the results when matched item is presented 2nd.

## Discussion

We presented items sequentially in order to control for competition between items for attention during encoding. The self-prioritization effect in WM was significant regardless of the serial position of the self-referential item, thus suggesting that this effect does not result from superior attentional capture during WM encoding. We also observed a main effect of serial order, in the shape of a primacy effect, but crucially, this effect did not interact with social association, but exerted an independent effect. Note that this observation is compatible with previous empirical and simulated data showing a primacy rather than recency effect in immediate free-recall under a fast rate of cue presentation (Davelaar et al., 2005). In sum, the results of Experiment 2 suggest that self-prioritization in WM is not driven by attentional or perceptual prioritization processes occurring at encoding. Combined with the fact that there are no stimulation differences during the probe/retrieval stage, this suggests that the self-bias results from the manner in which the items are being maintained in WM.

It is widely thought that WM maintenance involves directing attention internally to keep the to-be-maintained representations activated (Awh & Jonides, 2001; Barrouliet, Bernadin, & Camos, 2004; Oberauer, 2009; Chun, Golomb, & Turk-Browne, 2011; D'Esposito & Postle, 2015; Gazzaley & Nobre, 2012; Kiyonaga & Egner, 2013). Accordingly, previous studies have documented that when attention is directed towards a location stored in WM, this results in improved processing of external stimuli at that location (Awh, Jonides, & Reuter-Lorenz, 1998). In Experiment 3, we employ this logic to test the hypothesis that the self-reference effect in WM results from a bias of internal attention towards self-relevant stimulus locations during WM maintenance.

## Experiment 3

## **Method**

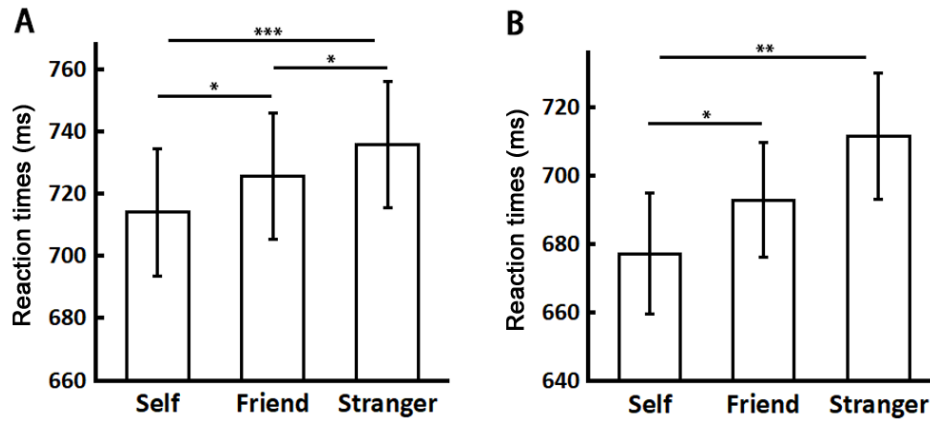
**Participants.** Twenty-five undergraduates from Southwest University (5 males; mean age = 19.62 years, range = 18–22) completed the experiment.

**Stimuli and Procedure.** The stimuli and procedure were identical to those in Experiment 1, except that, following the design of Awh et al. (1998), we added a dot-probe component to the WM task (see **Fig. 1C**). Following the presentation of the WM items, there was a variable interval of 3.3–3.7 seconds, after which two small dots were presented at one of the to-be-remembered locations (see inset in **Fig. 1C**). Participants had to judge whether the dots were horizontal or vertical, using the index and middle fingers of right hand on the keypad's "1" and "2" keys. The position-to-response assignment was counterbalanced across participants. The presentation of the dot probe was terminated by the key press or after 1500 ms, and was then followed by a blank screen for 2000 ms. The remaining WM probe part of this task was identical to Experiment 1.

The spatial WM task of Experiment 3 involved one factor with three levels (social association: Self, Friend and Stranger). Each trial type was presented 96 times, with 48 dot probe trials shown at each of the two item locations. For the location probes, there were 24 match trials for each of the two items and 48 non-match trials. Thus, there were altogether 288 trials, broken down into 6 blocks of 48 trials each (following 10 practice trials). The experiment lasted around 90 minutes.

## **Results**

The learning task and color-label probe data replicated the results of Experiment 1, documenting robust learning and maintenance of color-label associations (see Supplemental Material). Mean WM accuracy was 98% for the dot-probe judgment, 99% for the location probe, and 96% for the color-label probe on match trials. For the dot-probe RT data, a repeated-measures one-way ANOVA with 3 levels (Self-location, Friend-location and Stranger-location) showed a significant main effect ( $F(2, 48) = 14.47$ ,  $p < 0.001$ ,  $\eta^2 = 0.38$ , see **Fig. 4A**), with RT at the Self-location being significantly shorter than at the Friend-location ( $p < 0.05$ ) and the Stranger-location ( $p < 0.001$ ), and Friend-location responses being faster than Stranger-location ones ( $p < 0.05$ ) (see the **Table S3** in Supplemental Material for full descriptive statistics). For the WM RT data of location probe match trials, the main effect of social association was again significant ( $F(2, 48) = 7.36$ ,  $p < 0.01$ ,  $\eta^2 = 0.24$ , see **Fig. 4B**). Specifically, Self-match trials were faster than Friend-match trials ( $p < 0.05$ ) and Stranger-match trials ( $p < 0.01$ ), and Friend-match responses were marginally faster than Stranger-match ones ( $p = 0.07$ ).



**Fig. 4** Mean RTs ( $\pm$  SEM) for Self-match, Friend-match and Stranger-match conditions of dot-probe judgement (A) and the location probe (B) in Experiment 3. \*\*\* represents  $p < 0.001$ , \*\* represents  $p < 0.01$ , \* represents  $p < 0.05$ .

## Discussion

The results of Experiment 3 showed a significant self-prioritization effect on the dot-probe task during WM maintenance, with faster probe categorization at the location that corresponded to the self-related item held in WM. In line with prior studies of rehearsal effects in spatial WM (Awh et al., 1998; Awh & Jonides, 2001), we interpret these data as reflecting attentional prioritization of the self-referential item in WM, thus supporting our hypothesis that self-related information appears to attract greater internal attention than other-related representations during WM maintenance. In a final experiment we sought to evaluate whether this WM self-bias simply reflects participants' preference or whether it displays some automaticity. A classic way to probe the relative automaticity of a cognitive process is to create conditions where applying that process would produce sub-optimal task performance. Previous research has shown that rarely probed items normally get de-prioritized in WM (Gunseli, van Moorselaar, Meeter, & Olivers, 2015; Berryhill, Richmond, Shay, & Olson, 2012; Shimi, Nobre, Astle, & Scerif, 2014), but if the self-bias in WM is non-strategic, we would expect this effect to survive such a manipulation.

## Experiment 4

### Method

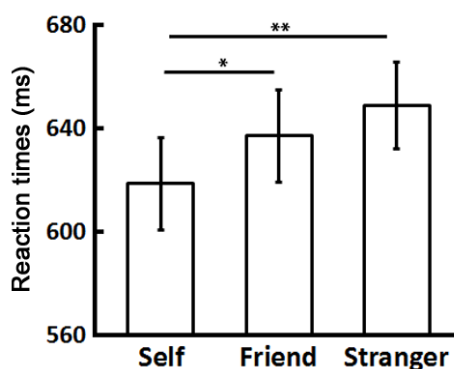
**Participants.** Twenty-three undergraduates from Southwest University (7 males; mean age = 20.59 years, range = 18–23) completed the experiment.

**Stimuli and Procedure.** The stimuli and procedure were identical to those in Experiment 1, except that in the spatial WM task we manipulated the probability of the self-matched item being probed. As before, the WM task involved one factor of three levels (social association: Self, Friend and Stranger), but the self-associated color was probed less than half as often (16.6%) as either of the other two colors (41.7% each). There were 384 trials total, with equal numbers of trials (128) for the different

color combinations (Self-Friend, etc.), broken down into 8 blocks of 48 trials each (following 10 practice trials). Importantly though, within each block, there were 4 self-matched trials, 10 friend-matched trials, 10 stranger-matched trials and 24 non-matched trials. Thus, the breakdown of self to friend and stranger probes was 2:5:5. The experiment lasted around 80 minutes.

## Results

The associative learning task data replicated the results of Experiment 1, documenting robust acquisition of color-label associations (see Supplemental Material). Mean WM accuracies for the location and color-label probes were 98% and 96%, respectively. A repeated measures one-way ANOVA on the RT data from match trials for the location probes revealed a significant main effect of social association ( $F(2, 44) = 8.02$ ,  $p < 0.01$ ,  $\eta^2 = 0.27$  see **Fig. 5**). Importantly, in spite of being rarely probed, the mean RT of Self-match trials was nevertheless significantly shorter than that of Friend-match trials ( $p < 0.05$ ) and Stranger-match trials ( $p < 0.01$ ), with Friend-match responses being marginally faster than stranger-match ones ( $p = 0.08$ ).



**Fig. 5** Mean RTs ( $\pm$  SEM) for Self-match, Friend-match and Stranger-match conditions of the location probe of spatial WM task in Experiment 4. \*\* represents  $p < 0.01$ , \* represents  $p < 0.05$ .

## Discussion

In sum, even though we rendered self-related probes a low-probability event, which has previously been shown to lead to a de-prioritization in WM (Gunseli et al., 2015; Berryhill et al., 2012; Shimi et al., 2014), the self-prioritization effect in WM was still significant. This result is consistent with a recent study showing a robust advantage for bottom-up prioritization of self-related stimuli under low probability (Sui et al., 2014). The present results suggest that biases favoring the representation of self-referential items in WM occur relatively automatically, as this prioritization persists in conditions where it represents a sub-optimal task strategy.

## General Discussion

We developed a novel experimental paradigm to test whether self-referential items are prioritized in WM. Across 4 experiments, results consistently showed a strong self-prioritization effect, with probes at locations associated with self-referential items being responded to faster than probes at locations associated with other social agents. Moreover, we observed a clear and reliable gradient of prioritization, with self-related stimuli being the most accessible, followed by friend-related and finally stranger-related stimuli. While self-bias effects are well established in the domains of exogenous attention (Brédart, Delchambre, & Laureys, 2006; Devue & Brédart, 2008), perceptual processing (Sui et al., 2012; Humphreys & Sui, 2015), and long-term memory (Kesebir & Oishi, 2010; Symons & Johnson, 1997), the present study is the first to document a self-prioritization effect in WM.

The current experiments also clearly speak to the locus of this self-bias effect. First, by adopting the associative learning design of Sui et al. (2012), we preempted the possibility that greater familiarity with self-related stimuli could influence the data. Second, by probing the remembered locations with a neutral, black circle cue (rather than with colors from the memory set), we equated visual input/bottom-up salience across conditions during WM retrieval. This ruled out the possibility that the WM advantage for self-related stimuli was driven by differential exogenous attention capture during retrieval. Third, Experiment 2 explicitly tested whether the WM self-bias was driven by differential encoding, specifically, by preferential attention to, or ordering of, the self-related item relative to the other items. While order-of-encoding did impact WM performance, this effect occurred in parallel, and did not interact, with the effect of self-prioritization. Thus, the latter effect most likely stems from a prioritized maintenance of the self-related item in WM. An interesting way of corroborating this conclusion further would be to develop a retro-cue version of the current task that focuses more exclusively on the WM maintenance stage (for review, see Souza & Oberauer, 2016).

Many current views on WM posit that items are being maintained by mechanisms of internally directed attention (Awh & Jonides, 2001; Barrouliet, Bernadin, & Camos, 2004; Oberauer, 2009; D'Esposito & Postle, 2015; Gazzaley & Nobre, 2012; Kiyonaga & Egner, 2013). By this logic, the self-related stimulus held in WM attracts internal attention in a similar fashion that an external self-related stimulus captures exogenous attention. In Experiment 3, we put this assumption to the test by adopting the design of Awh et al. (1998), testing for an attentional advantage for external probes presented at the self-related WM location during the maintenance interval. The results clearly supported the notion that the remembered location of the self-related item received more attention than locations of other-related items. Thus, in line with the assumption that equivalent mechanisms drive the allocation of internally and externally directed attention (Kiyonaga & Egner, 2013), we replicate an attentional bias for self-related items that had previously been demonstrated for exogenous attention in the domain of internal attention. A worthwhile addition to the present study in this regard might be the tracking of eye-movements, which could supply additional evidence for the preferred maintenance of the self-related item location over others (Williams, Pouget, Boucher, & Woodman, 2013).

Given that maintaining and evaluating information in WM is central to decision-making and cognitive control (Baddeley, 2003; D'Esposito & Postle, 2015), the present results suggest that a self-prioritization in WM could lead to undesirable social outcomes, such as promoting egocentric decision-making. Whether this is a serious concern depends in great part on the degree to which this self-bias is subject to strategic control, i.e., whether people are able to strategically override the prioritization of self-related stimuli in WM if that were advantageous. As an initial test of the automaticity of self-prioritization in WM, Experiment 4 pitted the self-bias against task statistics, by rendering probes of the self-related items less than half as likely as probes of other-related items. Previous studies have shown that, in the absence of self-bias, items with a low likelihood to be probed become deprioritized (Gunseli, van Moorselaar, Meeter, & Olivers, 2015; Berryhill, Richmond, Shay, & Olson, 2012; Shimi, Nobre, Astle, & Scerif, 2014). By contrast, the results of Experiment 4 showed that the self-bias in WM was robust to this probability manipulation. This suggests that self-prioritization in WM occurs automatically, recapitulating similar findings in perceptual processing (Sui et al., 2014), and highlights the importance of conducting future research to assess whether WM self-bias has negative real-life consequences; for instance, whether individual differences in WM self-prioritization are predictive of the degree to which people make selfish decisions.

In considering limitations of the present study, it should be noted that we cannot determine whether prioritization in WM of self-related items is reflected only in speeded access or may also be related to increased fidelity of self-related WM representations, because item precision was not a limiting factor in the current task. Future variants of this protocol could be designed to render WM performance more challenging, allowing one to assess WM accuracy and precision effects. Additionally, with respect to probing the automaticity of the self-bias in WM is, one could devise a stronger test than the manipulation we employed in Experiment 4, for instance, by probing the self-related item location even less frequently or by selectively rewarding performance for the other locations. It should also be noted that our inference in Experiment 4 is based on a *lack* of a frequency-driven change in performance, which presupposes that participants should be sensitive to such a probability manipulation in the first place. While we did not demonstrate this explicitly in the present study, the sensitivity to frequency changes of other-associated (but not self-associated) stimuli has been previously documented in the context of externally oriented attention in a highly similar protocol (Sui et al., 2014).

In summary, over four experiments, we document, for the first time, a robust prioritization of self-referential stimuli in WM. This effect appears to stem from an automatic attentional prioritization of self-related items during WM maintenance, and complements the attraction of self-related stimuli on exogenous attention demonstrated in previous research. The present findings expand our understanding of the impact of self-reference on high-level cognitive processing, and of the interaction between social salience, attention, and WM mechanisms, which may have profound impacts on social decision-making.

## Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

## Author Contributions

SY and AC initially conceived and designed the experiments, which were later improved and modified by discussions between AC, JS, YC, and TE. SY performed the experiments. The data were analyzed by SY, AC, and TE. The paper was written by SY, TE, AC, JS, and YC.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (61431013, 31771254), and the Fundamental Research Funds for the Central Universities (SWU1609106, SWU1709107).

## References

- Alexopoulos, T., Muller, D., Ric, F., & Marendaz, C. (2012). I, me, mine: Automatic attentional capture by self-related stimuli. *European Journal of Social Psychology*, 42, 770–779. doi:10.1002/ejsp.1882
- Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences*, 5, 119–126. doi:10.1016/S1364-6613(00)01593-X
- Awh, E., Jonides, J., & Reuter-Lorenz, P. A. (1998). Rehearsal in spatial working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 780–790. doi:10.1037/0096-1523.24.3.780
- Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience*, 4, 829–839. doi:10.1038/nrn1201
- Banaji, M. R., & Prentice, D. A. (1994). The self in social contexts. *Annual Review of Psychology*, 45, 297–332. doi:10.1146/annurev.ps.45.020194.001501
- Berryhill, M.E., Richmond, L.L., Shay, C.S., & Olson, I.R. (2012). Shifting attention among working memory representations: testing cue type, awareness, and strategic control. *Quarterly Journal of Experimental Psychology*, 65, 426–438.
- Blalock, L. D., & Clegg, B. A. (2010). Encoding and representation of simultaneous and sequential arrays in visuospatial working memory. *The Quarterly Journal of Experimental Psychology*, 63, 856–862. doi:10.1080/17470211003690680
- Brédart, S., Delchambre, M., & Laureys, S. (2006). One's own face is hard to ignore. *The Quarterly Journal of Experimental Psychology*, 59, 46–52. doi:10.1080/17470210500343678
- Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology*, 62, 73–101. doi:10.1146/annurev.psych.093008.100427

- Davelaar, E.J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H.J., Usher, M. (2005). The demise of short-term memory revisited: empirical and computational investigations of recency effects. *Psychological Review*, 112, 3-42.
- D'Esposito, M., & Postle, B. R. (2015). The cognitive neuroscience of working memory. *Annual Review of Psychology*, 66, 115–142.  
doi:10.1146/annurev-psych-010814-015031
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193–222.  
doi:10.1146/annurev.ne.18.030195.001205
- Devue, C., & Br édart, S. (2008). Attention to self-referential stimuli: Can I ignore my own face? *Acta Psychologica*, 128, 290–297. doi:10.1016/j.actpsy.2008.02.004
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175–191. doi:10.3758/bf03193146
- Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: bridging selective attention and working memory. *Trends in Cognitive Sciences*, 16, 129–135.  
doi:10.1016/j.tics.2011.11.014
- Griffin, I.C., Nobre, A.C. (2003). Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience*, 15, 1176-1194.
- Gunseli, E., van Moorselaar, D., Meeter, M., & Olivers, C. N. L. (2015). The reliability of retro-cues determines the fate of noncued visual working memory representations. *Psychonomic Bulletin & Review*, 22, 1334–1341.  
doi:10.3758/s13423-014-0796-x
- Harris, C. R., & Pashler, H. (2004). Attention and the processing of emotional words and names: Not so special after all. *Psychological Science*, 15, 171–178.  
doi:10.1111/j.0956-7976.2004.01503005.x
- Hay, D. C., Smyth, M. M., Hitch, G. J., & Horton, N. J. (2007). Serial position effects in short-term visual memory: A SIMPLE explanation? *Memory & Cognition*, 35, 176–190. doi:10.3758/bf03195953
- Humphreys, G. W., & Sui, J. (2015). The salient self: social saliency effects based on self-bias. *Journal of Cognitive Psychology*, 27, 129–140.  
doi:10.1080/20445911.2014.996156
- Kesebir, S., & Oishi, S. (2010). A spontaneous self-reference effect in memory: Why some birthdays are harder to remember than others. *Psychological Science*, 21, 1525–1531. doi:10.1177/0956797610383436
- Keyes, H., & Brady, N. (2010). Self-face recognition is characterized by “bilateral gain” and by faster, more accurate performance which persists when faces are inverted. *The Quarterly Journal of Experimental Psychology*, 63, 840–847.  
doi:10.1080/17470211003611264
- Kiyonaga, A., & Egner, T. (2013). Working memory as internal attention: Toward an integrative account of internal and external selection processes. *Psychonomic Bulletin & Review*, 20, 228–242. doi:10.3758/s13423-012-0359-y
- Liu, M., He, X., Rotstein, P., & Sui, J. (2016). Dynamically orienting your own face facilitates the automatic attraction of attention. *Cognitive neuroscience*, 7, 37–44.



doi:10.1080/17588928.2015.1044428

- Oberauer, K. (2009). Design for a Working Memory. In B.H. Ross (Ed.), *The Psychology of learning and motivation: Volume 51* (pp. 45–100). San Diego, CA: Elsevier Academic Press.
- Shimi, A., Nobre, A.C., Astle, D.E., & Scerif, G. (2014). Orienting attention within visual short-term memory: Development and mechanisms. *Child Development*, 85, 578-592.
- Souza, A.S., Oberauer, K. (2016). In search of the focus of attention in working memory: 13 years of the retro-cue effect. *Attention, Perception, & Psychophysics*, 78, 1839-1860.
- Sui, J., He, X., & Humphreys, G. W. (2012). Perceptual effects of social salience: evidence from self-prioritization effects on perceptual matching. *Journal of Experimental Psychology: Human Perception and Performance*, 38, 1105–1117. doi:10.1037/a0029792
- Sui, J., & Humphreys, G. W. (2015). The integrative self: How self-reference integrates perception and memory. *Trends in Cognitive Sciences*, 19, 719–728. doi:10.1016/j.tics.2015.08.015
- Sui, J., Sun, Y., Peng, K., & Humphreys, G. W. (2014). The automatic and the expected self: separating self- and familiarity biases effects by manipulating stimulus probability. *Attention, Perception, & Psychophysics*, 76, 1176–1184. doi:10.3758/s13414-014-0631-5
- Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: a meta-analysis. *Psychological bulletin*, 121, 371–394. doi:10.1037/0033-2909.121.3.371
- Williams, M., Pouget, P., Boucher, L., Woodman, G.F. (2013). Visual-spatial attention aids the maintenance of object representations in visual working memory. *Memory & Cognition*, 41, 698-715.